

DOI:10.19651/j.cnki.emt.2417354

面向复杂环境的改进 YOLOv5 安全帽检测算法^{*}

宋春宇 李寅中

(广西大学电气工程学院 南宁 530004)

摘要: 对施工工人的安全帽佩戴检测是保障人员安全的重要方法,但现有的安全帽检测大多为人工检测,不仅耗时费力且效率低下。且目前存在的算法在面对复杂的环境或者天气下,存在检测精度低等问题。针对这一现象,基于 YOLOv5s 算法提出一种改进的安全帽佩戴检测算法。首先,基于残差思想和大型可分离模块设计提出 SLSKA-POOL 模块,并在池化层使用,该模块可以使网络更加关注目标特征,进一步提高网络能力;其次,提出 CAKConv 卷积模块,该模块通过不规则的卷积操作高效的提取特征,以提高网络性能;最后,在主干添加 EMA 模块,聚合多尺度空间结构信息,建立长短依赖关系,以获得更好的性能。实验结果表明:改进的 YOLOv5 与原算法相比,检测精度提升 2.2%,mAP@0.5 提升了 3.6%,mAP@0.5:0.95 提升了 6.4%,实现了更准确高效的安全帽佩戴检测。

关键词: YOLOv5;安全帽检测;注意力机制;CAKConv;data augmentation

中图分类号: TN391 **文献标识码:** A **国家标准学科分类代码:** 520.6040

Improved YOLOv5 safety helmet detection algorithm for complex environments

Song Chunying Li Yinzhong

(School of Electrical Engineering, Guangxi University, Nanning 530004, China)

Abstract: Detecting the wearing of safety helmets by construction workers is an important method to ensure personnel safety. However, existing safety helmet detection methods are mostly manual, which are not only time-consuming and labor-intensive but also inefficient. Moreover, the existing algorithm has low detection accuracy in the face of complex environment or weather. In response to this phenomenon, an improved safety helmet wearing detection algorithm is proposed based on the YOLOv5s algorithm. Firstly, the SLSKA-POOL module is proposed based on the residual idea and large separable module design, and used in the pooling layer. This module can make the network pay more attention to the target features and further improve the network capability; secondly, the CAKConv convolutional module is proposed, which efficiently extracts features through irregular convolution operation to improve the network performance; finally, EMA modules are added to the backbone to aggregate multi-scale spatial structure information and establish short and short dependencies to achieve better performance. The experimental results show that: the improved YOLOv5 compared with the original algorithm, The detection accuracy increased by 2.2%, mAP@0.5 increased by 3.6%, and mAP@0.5:0.95 increased by 6.4%, realizing more accurate and efficient helmet wearing detection.

Keywords: YOLOv5; safety helmet detection; attention mechanism; CAKConv; data augmentation

0 引言

为了安全防护高空物体坠落对施工人员的伤害,关键措施是要求施工人员佩戴好安全帽。安全帽的使用能有

效减轻物体坠落对头部的冲击力,极大程度上确保施工人员的安全,减少头部受伤的可能性。当前大部分的施工现场,对工人佩戴安全帽情况的检查主要通过人工来进行,但这种方式效率较低。采用计算机视觉的手段,能有

收稿日期:2024-11-13

* 基金项目:国家自然科学基金(51767005)项目资助

效提升对检测工人佩戴安全帽的效率和准确性,因此,利用摄像头对施工现场的工人进行安全帽佩戴检测具有现实意义。

近年来,深度学习领域的飞速进展催生一系列的技术突破,引发人们广泛关注,其中卷积神经网络由于独特的卷积核机制,已经成为目标检测领域的热门选择。对于深度学习的目标检测算法领域,大致分为两类:双阶段检测算法与单阶段检测算法。双阶段检测算法主要以 R-CNN^[1]、Fast R-CNN^[2]、Faster R-CNN^[3]、Mask R-CNN^[4] 等为代表,单阶段检测算法以 YOLO(you only look once)^[5-8] 系列、Retina-Net^[9]、SSD(single shot multiBox detector)^[10] 为代表的目标检测算法,2015年,YOLO算法被提出,因其优异的检测性能,YOLO算法得到快速发展。

丁田等^[11]针对安全帽检测问题,首先在算法中加入通道注意力机制(efficient channel attention,ECA)来提高特张提取能力;其次使用损失函数 CIOU(complete intersection over union)计算目标损失,来加速收敛;最后重新构建安全帽数据集来验证改进的检测算法。许迪等^[12]首先基于空洞卷积和残差思想设计一个新的模块来增加感受野;其次设计一个轻量的特张金字塔结构来提取和融合特征;他们最后引入新的损失函数来优化边界框的回归。刘昶成等^[13]提出一种新的轻量化算法:他们首先在主干加入深度可分离卷积来降低网络参数量;其次又添加坐标注意力机制(coordinate attention,CA)以提升检测精度;最后引入 EIOU 损失函数。李永辉等^[14]将 ECA 注意力机制与卷积模块融合,其促进全局特征信息的融合;最后使用高效损失函数(efficient intersection over union,EIOU)来优化模型的损失值。虽然上述改进算法在检测中表现良好,但其检测精度仍有提升空间。

虽然上述的算法研究在安全帽检测中取得一定的效果,但在面对复杂的检测环境检测图中小目标,算法依然存在检测准确率低、漏检、误检的情况。针对以上问题,本文使用的基准模型为 YOLOv5s。使用改进的 YOLOv5s 来对安全帽佩戴进行检测。在原始的算法上,首先优化算法的池化模块,使网络更加关注目标特征,使模型更好的检测不同复杂环境图像中的安全帽特征,以提高网络检测能力。其次,提出 CAKConv 卷积模块,模块通过不规则的卷积操作高效的提取目标特征,使得模型可以更好的检测因拍摄角度不同导致检测目标大小变化或者检测目标被部分遮挡的情况,以提高网络性能。最后,在主干网络添加 EMA 注意力模块,使网络聚合多尺度空间结构信息,建立长短期依赖关系。

1 YOLOv5 网络介绍

YOLOv5 作为 YOLO 家族的一员,它不仅继承 YOLO 系列高检测精度和快速推理的优势,同时还优化模型的紧凑性,拥有更小巧的模型体积。YOLOv5 有多个不

同的版本,这些版本仅在网络的深度和宽度不同。

YOLOv5 网络由 4 部分组成:输入端、主干网络、颈部以及输出端。为了让网络有更好的泛化性算法在输入阶段将 4 幅照片随机放缩、剪切的方式进行拼接,来丰富训练数据集。最后将图像缩放至统一的尺寸以便后续的训练。YOLOv5 的主干网络,作为特征提取的核心结构,由 CBS、C3、SPPF 模块组成。CBS 模块由二维卷积 Conv,批量归一化处理和 SiLU 激活函数串联而成。C3 模块包含 3 个卷积核,主要用途为提高网络提取图像信息的能力。SPPF 模块采用多个并行的池化模块来融合不同感受野的特征信息。颈部部分由特征金字塔网络(feature pyramid network,FPN)和路径聚合网络(path aggregation network,PAN)^[15]结构组成。FPN 将深层的特征通过自顶向下的上采样方式传输融合到浅层特征中,从而实现强语义特征的传输;而 PAN 则自底向上传递浅层特征的强定位信息。FPN 和 PAN 一起将网络获取的语义信息和位置信息高效融合在一起,输出到预测部分。预测端包括边界框损失函数及非极大值抑制算法(non maximum suppression,NMS),它采用 CIOU 作为损失函数对目标进行回归修正,有效提升预测框回归的速度以及精度。同时使用加权非极大值抑制来消除大量冗余框来进一步提高对目标的识别准确度。算法结构如图 1 所示。

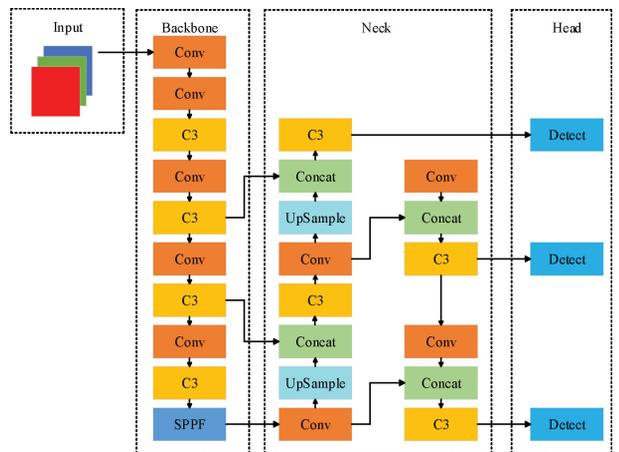


图 1 YOLOv5s 网络结构

Fig.1 YOLOv5s network structure

2 改进 YOLOv5s 模型

2.1 改进网络池化结构

安全帽图像数据中存在着许多小目标来检测,但对于这些小目标来说,它们仅仅占据着图像的一小部分,容易受到其他非检测对象和背景的干扰。为了进一步使模型更为关注安全帽的形状特征,提出了 SLSKA-POOL 模块。该模块可以使模型能够在特征表示中编码更多的形状信息和更少的纹理。

大可分离核注意力模块 (large separable kernel attention, LSKA)^[16] 结构如图 2 所示。对于一个输入特征图 $F \in R^{C \times H \times W}$, 其中 C 是输入通道数, H 和 W 分别表示特征图的高度和宽度。

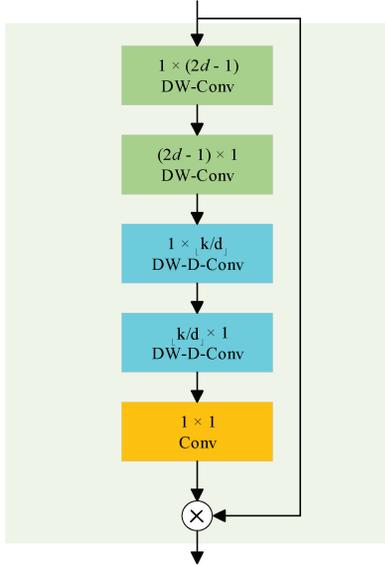


图 2 LSKA 模块
Fig. 2 LSKA module

通过将深度卷积和深度扩展卷积的二维权值核分割成两个级联一维可分离权值核, 可以得到 LSKA。LSKA 的输出为:

$$ZC = \sum_{H,W} W_{(2d-1) \times 1}^C * (\sum_{H,W} W_{(2d-1) \times 1}^C * F^C) \quad (1)$$

$$Z^C = \sum_{H,W} W_{\lfloor \frac{k}{d} \rfloor \times 1}^C * (\sum_{H,W} W_{1 \times \lfloor \frac{k}{d} \rfloor}^C * Z^C) \quad (2)$$

$$A^C = W_{1 \times 1} * Z^C \quad (3)$$

$$\bar{F}^C = A^C \otimes F^C \quad (4)$$

其中, d 是膨胀速率, $*$ 表示卷积, \otimes 表示 Hadamard 积, $W_{k \times k}$ 表示 $K \times K$ 大小卷积核。 Z^C 表示级联后的深度卷积的输出, 该输出捕获局部空间信息。 $\lfloor \cdot \rfloor$ 代表对变量向下取整。扩展的深度卷积负责捕获深度卷积的输出 Z^C 的全局空间信息。 A^C 为得到的注意力图。大内核是捕获全局信息的一种方法。LSKAAtten 方法可以有效地捕获图像中的长期依赖性。

由于实际的生活生产中, 安全帽佩戴检测时的背景环境是复杂多变的, 检测时会受到当时天气的影响而产生不

同的检测结果, 因此本文提出 SLSKA-POOL 模块, 优化算法的池化层, 使得网络在聚合不同尺度的特征时加强对目标特征的关注, 降低背景信息的影响。SLSKA-POOL 模块结构如图 3 所示。

2.2 CAKConv 模块

在标准卷积运算中存在两个固有的缺陷。一方面, 卷积操作局限于一个局部窗口, 不能从其他位置捕获信息, 其采样形状是固定的。另一方面, 卷积核的大小是一个固定的正方形。但目标的形状大小在不同的图像上是不同的。具有固定形状的卷积核不能适应不断变化的目标。针对上述问题, 本文提出 CAKConv 卷积模块, 该模块为网络提供任意采样形状, 对网络开销和性能之间的权衡提供更丰富的选择。

在可变核卷积 (alterable kernel convolution, AKConv)^[17] 中, 通过一种新的坐标生成算法, 生成卷积核 P_n 的初始采样坐标。它首先将采样网格作为常规采样网格生成, 然后为剩余的采样点创建不规则网格, 最后, 将它们缝合起来, 生成整体采样网格。为了适应所使用的卷积的尺寸, 算法设置左上角 $(0, 0)$ 点作为采样原点。在定义了不规则卷积的初始坐标 P_n 后, 在位置 P_0 处对应的卷积操作可以定义为:

$$Conv(P_0) = \sum w \times (P_0 + P_n) \quad (5)$$

其中, w 表示卷积参数。

为适应目标的变化, 引入偏移量来调整每个位置上的样本形状, 核的偏移量首先通过卷积运算得到, 然后将偏移坐标和原始坐标 $(P_0 + P_n)$ 相加, 得到修改后的坐标。最后通过插值和重采样得到相应位置的特征。为提取出不规则卷积核的采样位置所对应的特征, 可以将特征堆叠在行或列上, 以使用列卷积或行卷积来提取与不规则采样形状相对应的特征。因此所提模块可以完美地完成不规则卷积特征提取过程。通过不规则卷积完成特征提取过程, 根据偏移量灵活调整样本形状, 为卷积采样形状带来更多的探索选择。AKConv 卷积模块结构如图 4 所示。

对于在复杂场景下的安全帽佩戴检测可能导致图像中安全帽大小的变化或者安全帽被部分遮挡等情况, 本文因此提出 CAKConv 模块, 该模块是结合 C3 和 AKConv 所构建的一个新的卷积模块, 结构图如图 5 所示。模块使得网络在面对不同的目标形状和大小时, 可以灵活的进行卷积采样, 很好的解决因复杂环境导致的安全帽形状变化的问题。

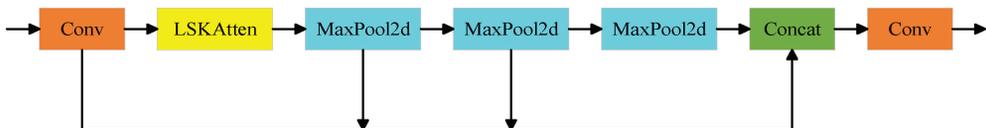


图 3 SLSKA-POOL 模块结构
Fig. 3 The SLSKA-POOL module structure

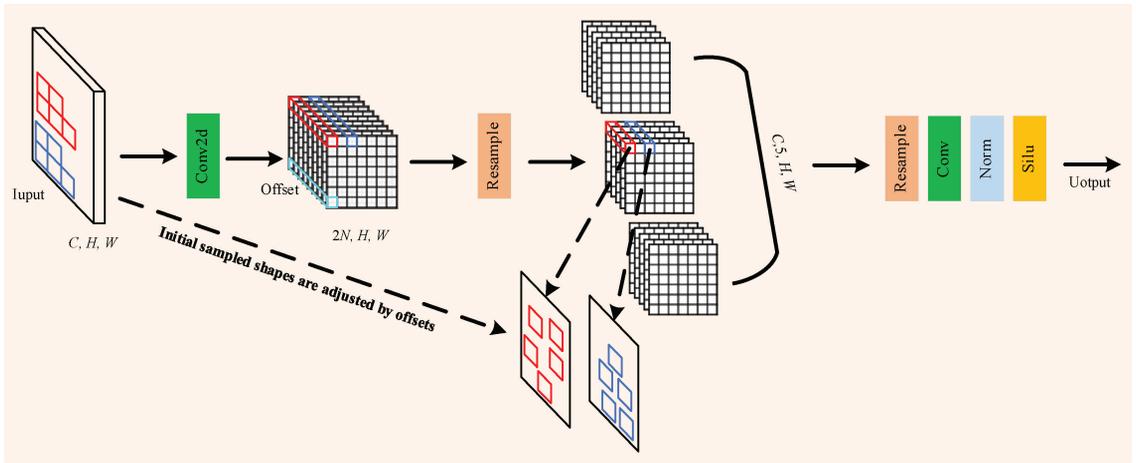


图 4 AKConv 结构

Fig. 4 The AKConv structure

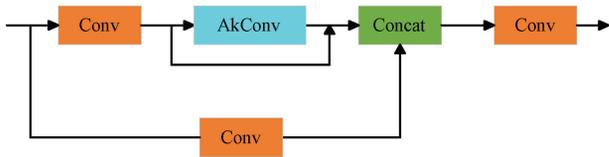


图 5 CAKConv 结构

Fig. 5 The CAKConv structure

2.3 EMA 高效多尺度注意力机制

高效注意力机制 (efficient multi-scale attention, EMA)^[18] 利用全局平均池操作来建模跨通道信息。其结构如图 6 所示。对于特征图 $X = R^{C \times H \times W}$, EMA 将在通道维度方向上将 X 划分为 G 个子特征, 以学习不同的语义信息。其中, 模块的分组方式可以表示为 $X = [X_0, X_1, \dots, X_{G-1}]$, $X_i = R^{C//G \times H \times W}$ 。

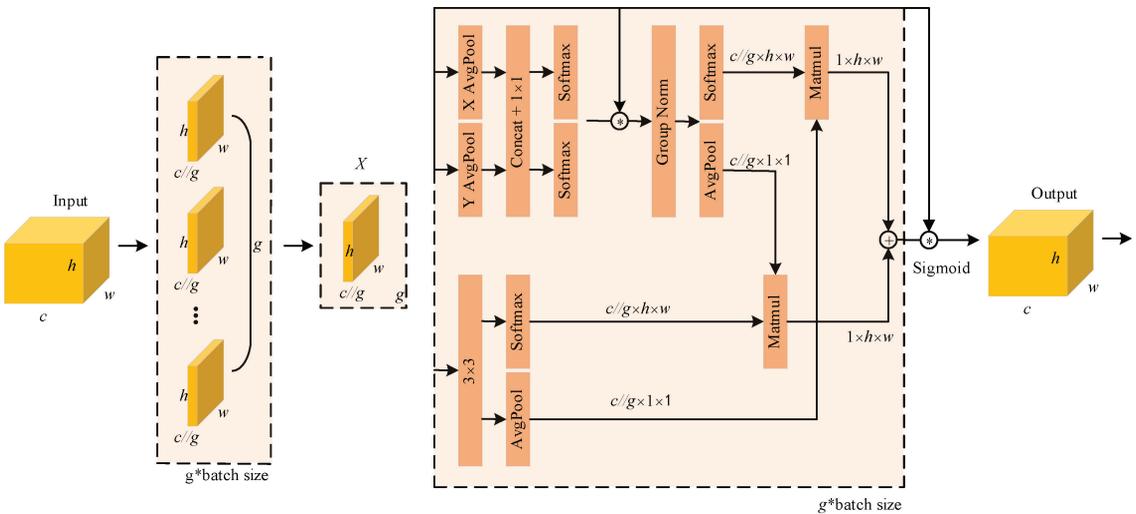


图 6 EMA 结构

Fig. 6 EMA structure

神经元较大的局部感受野能够收集多尺度的空间信息。因此为了聚合多尺度的空间结构信息, EMA 利用 3 条平行路径来提取分组特征图的注意力权重描述符。其中两条平行路由是 1×1 分支, 第 3 条路由是 3×3 分支。为了捕获通道之间的依赖关系对通道方向上的交叉通道信息交互进行了建模。在 1×1 分支中, 有两个一维全局平均池化操作分别沿着两个空间方向对通道进行编码, 另一方面, 3×3 分支通过 3×3 卷积捕获局部跨通道交互, 以扩

大特征空间。EMA 不仅对信道间信息进行编码, 以调整不同信道的重要性, 并将精确的空间结构信息保存到信道中。最后利用二维全局平均池对 1×1 分支输出中的全局空间信息进行编码, 在通道特征的联合激活机制之前, 将最小分支的输出直接转换为相应的维度形状, 即 $R_1^{1 \times C//G \times R_3^{C//G \times H \times W}}$ 。二维全局池化操作的表述为:

$$Z_c = \frac{1}{H \times W} \sum_j \sum_i x_c(i, j) \quad (6)$$

将上述并行处理的输出与矩阵点积运算相乘后,得到了第1个空间注意力图。利用二维全局平均池对 3×3 分支中的全局空间信息进行编码,在通道特征的联合激活机制之前,将 1×1 分支直接转换为相应的维数形状,即 $R_3^{1 \times C // G} \times R_1^{C // G \times HW}$ 。然后,得到了保留了整个精确空间位置信息的第2张空间注意力图。最后,将每一组内的输出特征图计算为所生成的两个空间注意力权重值的聚合,然后是一个Sigmoid函数。它捕获像素级的成对关系,并突出显示所有像素的全局上下文。

由于受到复杂环境影响,在对目标进行检测时其背景通常会变得复杂多变,算法会容易出现误检、漏检等情况。通过在算法中引入EMA注意力机制模块,减少无关特征的干扰,可以解决算法在安全帽佩戴检测中出现误检、漏检等问题。EMA的引入使得模型能够在不同尺度的特征空间中进行动态的加权,从而捕捉更丰富的信息。使得模型不仅关注当前尺度的细节,还能够捕捉到较大尺度的全局信息。这样可以帮助模型更好地理解复杂的输入图像。

3 实验结果与分析

3.1 数据集

数据集的质量好坏对于模型的训练有着很大的影响。因此本文使用包含大量的安全帽检测图像的数据集 safety helmet wearing 数据集,该数据集共包含 7 581 张图片,涵盖了不同场景下、不同光照和不同遮挡的情况。本文按照 8:2 将数据集划分为训练集和验证集。

虽然原始的数据集有着很多安全帽图像,但为了提高数据集的复杂性和多样性,模拟实际生产施工时施工人员所处的天气环境情况。对所划分的训练集数据的部分图像进行数据增强,模拟实际的天气情况,包括:黑夜、大雾、下雨和下雪情况,同时也模拟了实际摄像可能会出现的一些情况,如:图片的垂直翻转、模糊图像和图像旋转的情况。通过使用上述的图像增强方法,模拟出了实际安全帽检测的复杂环境,极大的丰富了安全帽数据集的复杂性和多样性。

3.2 实验环境和训练参数

研究使用可以提供稳定运行环境的 Windows11 操作系统,使用 PyTorch 深度学习框架进行相关实验。在 NVIDIA GeForce RTX 4080 GPU 上进行训练。

对于训练,使用预训练权重有助于模型更快收敛,因此本实验采用预训练权重参数作为模型训练的初始参数设置。一些实验参数设置如表 1 所示。

3.3 评价指标

在本文中,选用平均精度 AP 和均值平均精度 mAP 作为核心评估指标,用于量化和比较改进后模型在目标检测任务中的性能表现。

召回率(recall, R)和精确率(precision, P)的公式如式

(7)和式(8)所示。

表 1 训练参数

Table 1 Training parameters

| 参数 | 参数值 |
|--------|---------|
| 输入图像大小 | 640×640 |
| 训练次数 | 200 |
| 批处理量 | 16 |
| 初始学习率 | 0.01 |
| 权重衰减系数 | 0.0005 |

$$P = \frac{TP}{TP + FP} \quad (7)$$

$$R = \frac{TP}{TP + FN} \quad (8)$$

其中,TP 代表了那些被模型正确地识别为正类别的样本数量,即被正确的预测为佩戴安全帽的样本;FP 则指代被模型错误地标记为正类别的负样本数目,即未佩戴安全帽的人被预测为佩戴了帽子的人;FN 表示那些实际上属于正类别,但却未能被模型检测到或正确分类的样本总量,即有安全帽的样本被判别为未佩戴安全帽的样本。

平均精度(AP)表示在不同的检测阈值下,模型精度的平均值。平均精度均值(mAP)则是对多个类别的

平均精度进行平均,以提供一个综合的性能评价。AP 和 mAP 计算公式分别为式(9)和式(10)。

$$AP = \int_0^1 P(R) dR \quad (9)$$

$$mAP = \frac{\sum_{i=1}^K AP_i}{K} \quad (10)$$

其中,K 表示检测目标类别数量。

3.4 消融实验

为了表现在 YOLOv5 架构中引入的各项改进模块的实际贡献,开展了一系列的消融实验。采用 YOLOv5s 作为基线模型,依次添加 EMA 模块、CAKConv 和 SLSKA-POOL 模块。将精度、召回率、mAP@0.5 等作为消融研究的指标,以验证每个模块的有效性。

消融实验旨在逐一评估每个改进模块对模型整体性能提升的作用。首先在算法的骨干网络中添加 EMA 模块;然后将 CAKConv 模块替换网络部分的 C3 模块,最后将提出的 SLSKA-POOL 模块替换进网络池化层中。从表中可以看到,替换 CAKConv 后,相比于原始的 YOLOv5 算法的 mAP@0.5 和 mAP@0.5:0.95 值,模型分别提升了 3.1% 和 5.4%。融合 SLSKA-POOL 后,模型分别提升了 3% 和 5.6%。添加 EMA 后,模型分别提升了 3% 和 5.1%。当结合所有改进时,算法的 mAP@0.5 和 mAP@0.5:0.95 值分别提高 3.2% 和 5%。最后加入数据增强后,mAP@0.5 和 mAP@0.5:0.95 值分别提升了 3.6% 和

6.4%。表 2 为消融实验的结果。

表 2 消融实验结果

Table 2 Results of the ablation experiments

| 模型 | P | R | $mAP@0.5$ | $mAP@0.5:0.95$ |
|-------------|-------|-------|-----------|----------------|
| YOLOv5 | 0.912 | 0.843 | 0.905 | 0.534 |
| +CAKConv | 0.931 | 0.882 | 0.936 | 0.588 |
| +SLSKA-POOL | 0.929 | 0.886 | 0.935 | 0.590 |
| +EMA | 0.932 | 0.884 | 0.935 | 0.585 |
| All | 0.932 | 0.889 | 0.937 | 0.584 |
| All+数据增强 | 0.934 | 0.894 | 0.941 | 0.598 |

3.5 对比实验

为了深入证实所提改进算法的性能,将改进算法与主流的目标检测算法进行对比评测。包括的算法有 SSD、YOLOv3、YOLOv5s、YOLOv5n、YOLOv8n 和 YOLOv10 n,从类别的平均精度及其均值、模型大小来进行比较,来证明改进算法的有效性。对比结果如表 3 所示。

表 3 主流算法对比

Table 3 Mainstream algorithm comparison

| 算法 | $AP(\text{hat})$ | $AP(\text{person})$ | $mAP@0.5$ | MB |
|--------------------------|------------------|---------------------|-----------|------|
| SSD | 0.813 | 0.527 | 0.670 | 91.1 |
| YOLOv3 | 0.699 | 0.886 | 0.789 | 16.6 |
| YOLOv5s | 0.900 | 0.918 | 0.909 | 13.7 |
| YOLOv5n | 0.897 | 0.915 | 0.906 | 3.6 |
| YOLOv8n ^[19] | 0.883 | 0.922 | 0.903 | 5.95 |
| YOLOv10n ^[20] | 0.903 | 0.918 | 0.924 | 5.48 |
| 本文 | 0.930 | 0.933 | 0.941 | 15.3 |

从表 3 可以看出,这些对比的主流算法在检测类别‘hat’时的平均精度大多低于 90%,与改进算法相比有较大差距;虽然它们在检测‘person’类别的平均精度有着不错的表现,但均低于改进算法。综合来说,本文所提改进算法较其他对比算法在评价指标上有着显著提升,其平均精度均值达到了 0.941,表明改进模型具有很好的检测效果。

3.6 检测结果分析

为了全面评估本文提出的 ALE-YOLOv5s 算法在面对复杂环境中检测安全帽的性能表现,本文将分别使用未改进算法和改进算法在验证集中进行预测与对比分析。图 7 直观的呈现了 ALE-YOLOv5s 算法与原始 YOLOv5s 模型在不同工作场景环境下对安全帽佩戴识别效果的对比,其中左侧图像展示的是基础模型 YOLOv5s 的安全帽识别结果,右侧则对应改进后 ALE-YOLOv5s 模型的检测结果。

图 7(a)为模拟大雾天气情况下的检测场景,未改进算

法的检测效果图为左侧图,改进算法的检测效果图为右侧图。可以看到,对于在大雾中检测时,未改进算法漏检了图左侧的目标;图 7(b)为在密集场景下对目标的检测结果,可以看到未改进算法存在误检的情况;图 7(c)为模拟存在模糊照片时对目标检测的场景,同样的,未改进算法存在漏检的问题;图 7(d)为在下雨检测的场景,这时未改进算法出现误检的情况;图 7(e)为模拟拍摄角度旋转的情况,左侧的未改进算法检测图存在着漏检的问题。图 7(f)为下雪检测场景,未改进算法出现了误检的情况;图 7(g)为对小目标检测场景,位于上方的图为未改进算法检测结果图,可以看到图中的上方存在着严重的漏检问题。而改进算法在面对这类小目标检测时有着不错的检测效果。由上面的检测结果,证明了改进后的模型在面对复杂环境和不同天气时具有较好的鲁棒性。综上可知,ALE-YOLOv5s 算法在面对复杂环境时检测效果表现较好,能够显著提升算法的检测性能。



(a) 大雾天气检测场景
(a) Fog weather detection scene



(b) 密集检测场景
(b) Intensive detection scenario



(c) 模糊照片检测场景
(c) Fuzzy photo detection scene



(d) 下雨天检测场景
(d) Detection scene in rainy days



(e) 旋转图片检测场景
(e) Rotate the picture for the detection scene



(f) 下雪检测场景
(f) Snow detection scene



(g) 小目标检测场景
(g) Small target detection scenario

图7 不同场景下的模型检测对比

Fig. 7 Model detection and comparison in different scenarios

4 结 论

对当前复杂环境下检测工人佩戴安全帽的算法存在的检测效能不足、出现漏检误检等问题,本研究提出一种改进YOLOv5s算法,旨在提升算法在面对密集目标或在复杂环境中对安全帽识别精度。算法的改进包括提出CAKConv模块、引入EMA注意机制以及在池化层中提出SLSKA-POOL模块的3个改进点。在使用了数据增强的数据集SHWD中进行实验。实验结果可知,改进后的模型在准确率、召回率、 $mAP_{0.5}$ 和 $mAP_{0.5:0.95}$ 分别提高了2.2%、5.1%、3.6%和6.4%。

实验结果说明了提出的ALE-YOLOv5s算法的有效性。特别是改进算法在复杂环境中对小尺寸的安全帽检测、密集目标检测有更优异识别精度。改进算法在识别工人是否佩戴安全帽中表现出了更好的效果。

参考文献

- [1] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [2] GIRSHICK R. Fast R-CNN[C]. IEEE International Conference on Computer Vision, 2015: 1440-1448.
- [3] REN SH Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [4] HE K M, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2020, 42(2): 386-397.
- [5] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2016: 779-788.
- [6] REDMON J, FARHADI A. YOLO9000: Better, faster, stronger [C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition. Honolulu: IEEE, 2017.
- [7] REDMON J, FARHADI A. YOLOv3: An incremental improvement [J]. Computer Vision and Pattern Recognition, 2018(8): 1-6.
- [8] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[J]. ArXiv preprint arXiv:2004.10934, 2020.
- [9] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection [C]. IEEE International Conference on Computer Vision, 2017: 2980-2988.
- [10] LIU W, ANGELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]. 2016 European Conference on Computer Vision (ECCV), Amsterdam, The Netherlands, 2016: 21-37.
- [11] 丁田, 陈向阳, 周强, 等. 基于改进YOLOX的安全帽佩戴实时检测[J]. 电子测量技术, 2022, 45(17): 72-78.
DING T, CHEN X Y, ZHOU Q, et al. Real-time detection of helmet wearing based on improved YOLOX [J]. Electronic Measurement Technology, 2022, 45(17): 72-78.
- [12] 许迪, 张淑卿, 葛超. 面向复杂环境的YOLOv8安全装备检测[J]. 电子测量技术, 2024, 47(7): 121-129.
XU D, ZHANG SH Q, GE CH. YOLOv8 security equipment detection for complex environments [J]. Electronic Measurement Technology, 2024, 47(7):

- 121-129.
- [13] 刘昶成,邵文权,李玲陶. 基于改进 YOLOv5s 的电力作业人员安全帽检测算法研究[J]. 国外电子测量技术,2024,43(2):34-42.
- LIU CH CH,SHAO W Q,LI L T. Research on safety helmet detection algorithm for power operators based on improved YOLOv5s [J]. Foreign Electronic Measurement Technology, 2024,43(2): 34-42.
- [14] 李永辉,袁亮,何丽,等. 化工园区场景下安全帽佩戴状态检测算法研究[J]. 计算机工程与应用,2024,60(12):111-117.
- LI Y H,YUAN L,HE L,et al. Research on the safety helmet wearing detection algorithm in chemical industry park scenarios[J]. Computer Engineering and Applications,2024,60(12):111-117.
- [15] LIU SH, QI L, QIN H F, et al. Path aggregation network for instance segmentation [C]. IEEE Conference on Computer Vision and Pattern Recognition,2018:8759-8768.
- [16] KIN W L, LAI M B, et al. Large separable kernel attention;Rethinking the large kernel attention design in CNN[C]. IEEE Conference on Computer Vision and Pattern Recognition(CVPR),2023.
- [17] ZHANG X,SONG Y Z,SONG T T,et al. AKConv: Convolutional kernel with arbitrary sampled shapes and arbitrary number of parameters [J]. ArXiv preprint arXiv:2311.11587, 2023.
- [18] OUYANG D L, HE S, ZHAN G ZH, et al. Efficient multi-scale attention module with cross-spatial learning[C]. ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing(ICASSP). IEEE, 2023:1-5.
- [19] VARGHESE R, SAMBATH M. YOLOv8: A novel object detection algorithm with enhanced performance and robustness[C]. 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems(ADICS). IEEE, 2024: 1-6.
- [20] WANG AO,CHEN H,LIU L H,et al. YOLOv10: Real-time end-to-end object detection[J]. Neural Information Processing Systems, 2025, 37: 107984-108011.

作者简介

宋春宁(通信作者),副教授,硕士生导师,主要研究方向为人工智能、FPGA、电力系统、电池。

E-mail:scn206@gxu.edu.cn

李寅中,硕士研究生,主要研究方向为图像处理。

E-mail:1002216619@qq.com